**Indian Nutrient Databank (INDB)**

This code generates a food composition database for India called the ‘Indian Nutrient Databank (INDB).’ INDB includes raw foods and standard recipes commonly consumed in India.

The following abbreviations are used in this README file.

ASC = Art & Science of Cooking

BFP = Basic Food Preparation

FAO = Food and Agriculture Organization of the United Nations

FCT = Food Composition Table

ICMR = Indian Council of Medical Research

IFCT = Indian Food Composition Table

NIN = National Institute of Nutrition

NRF = Nutrient Retention Factors

OSR = Open Source Recipes

USDA = US Department of Agriculture

**There are 7 data files:**

1. Excel file called ‘NIN\_fct.xlsx'

This contains nutrient data (amount of nutrient per 100g food item except in the case of alcoholic beverages which are presented per 100ml) for 897 raw food items. It is compiled from the ICMR-NIN IFCT 2004 and IFCT 2017. When duplicate food items were in the 2004 and 2017 editions, the more recent (2017) was retained.

NA indicates the value was not available in IFCT 2004 or 2017. This was the case for many protein foods for which carbohydrates, free sugars, and fibre were not measured. For commonly consumed foods, a secondary source (see UK\_fct file below) was consulted for missing nutrient values. In these instances, a trace value for a nutrient is represented by Tr and where a nutrient is present in significant quantities, but there is no reliable information on the amount, the value is represented by N.

Nutrients include: energy (kJ), energy (kcal), carbohydrates (g), protein (g), fat (g), free sugars (g), fibre (g), saturated fatty acid (mg), monounsaturated fatty acid (mg), polyunsaturated fatty acid (mg), cholesterol (mg), calcium (mg), phosphorus (mg), magnesium (mg), sodium (mg), potassium (mg), iron (mg), copper (mg), selenium (µg), chromium (mg), manganese (mg), molybdenum (mg), zinc (mg), vitamin A (µg), vitamin E (mg), vitamin D2 (µg), vitamin D3 (µg), vitamin K1 (µg), vitamin K2 (µg), folate (µg), thiamine (mg), riboflavin (mg), niacin (mg), pantothenic acid (mg), vitamin B6 (mg), biotin (µg), folic acid (µg), vitamin C (mg), and carotenoids (µg).

2. Excel file called ‘UK\_fct.xlsx'

This contains nutrient data (amount of nutrient per 100g food item except in the case of alcoholic beverages which are presented per 100ml) for 144 ingredients that are found in Indian recipes included in the INDB but were not in the NIN\_fct file. These data are from the 2021 version of the UK ‘McCance and Widdowson's composition of foods integrated dataset’ available from: <https://www.gov.uk/government/publications/composition-of-foods-integrated-dataset-cofid>.

A trace value for a nutrient is represented by Tr. Where a nutrient is present in significant quantities, but there is no reliable information on the amount, the value is represented by N.

Nutrients include: energy (kJ), energy (kcal), carbohydrates (g), protein (g), fat (g), free sugars (g), fibre (g), saturated fatty acid (mg), monounsaturated fatty acid (mg), polyunsaturated fatty acid (mg), cholesterol (mg), calcium (mg), phosphorus (mg), magnesium (mg), sodium (mg), potassium (mg), iron (mg), copper (mg), selenium (µg), chromium (mg), manganese (mg), molybdenum (mg), zinc (mg), vitamin A (µg), vitamin E (mg), vitamin D2 (µg), vitamin D3 (µg), vitamin K1 (µg), vitamin K2 (µg), folate (µg), thiamine (mg), riboflavin (mg), niacin (mg), pantothenic acid (mg), vitamin B6 (mg), biotin (µg), folic acid (µg), vitamin C (mg), and carotenoids (µg).

3. Excel file called ‘US\_fct.xlsx'

This contains nutrient data (amount of nutrient per 100g food item except in the case of alcoholic beverages which are presented per 100ml) for 54 ingredients that are found in Indian recipes included in the INDB but were not in the NIN\_fct file or the UK dataset mentioned in #2. These data are from the US Department of Agriculture’s FoodData Central available from: <https://fdc.nal.usda.gov/>. The prefix 'US' was added before the original USDA code to identify the source.

Where there was no information on the amount of a nutrient the value it is represented by NA.

Nutrients include: energy (kJ), energy (kcal), carbohydrates (g), protein (g), fat (g), free sugars (g), fibre (g), saturated fatty acid (mg), monounsaturated fatty acid (mg), polyunsaturated fatty acid (mg), cholesterol (mg), calcium (mg), phosphorus (mg), magnesium (mg), sodium (mg), potassium (mg), iron (mg), copper (mg), selenium (µg), chromium (mg), manganese (mg), molybdenum (mg), zinc (mg), vitamin A (µg), vitamin E (mg), vitamin D2 (µg), vitamin D3 (µg), vitamin K1 (µg), vitamin K2 (µg), folate (µg), thiamine (mg), riboflavin (mg), niacin (mg), pantothenic acid (mg), vitamin B6 (mg), biotin (µg), folic acid (µg), vitamin C (mg), and carotenoids (µg).

4. Excel file called ‘recipes.xlsx'

This contains the ingredients (name, amount, amount unit, and food code linking it to the NIN\_fct, UK\_fct, or US\_fct) of recipes. Recipes are from the following sources:

• 490 recipes from: Khanna K, Gupta S, Seth R et al. The Art & Science of Cooking. Fifth Edition. New Delhi: Elite Publishing, 2007.

• 378 recipes from: Raina U, Kashyap S, Narula V et al. Basic Food Preparation: A Complete Manual. Fourth Edition. New Delhi: Orient BlackSwan, 2011.

• 148 recipes from websites. See information file ‘recipe\_links’.

The food code naming conventions are: • ‘ASC’ as prefix for recipes from ‘The Art & Science of Cooking’ • ‘BFP’ as prefix for recipes from ‘Basic Food Preparation’ • ‘OSR’ as a prefix for recipes from websites

The recipe names have been updated to include the English name and local or common names of the recipes.

5. Excel file called ‘recipes\_names.xlsx'

This contains the common names of recipes from the dataset mentioned in #4.

6. Excel file called ‘recipes\_servingsize.xlsx'

This contains the serving sizes for recipes from the dataset mentioned in #4.

7. Excel file called ‘USDA\_nrf. xlsx’

This contains the information on nutrient retention factors from the USDA Table of Nutrient Retention Factors, Release 6 (2007): https://agdatacommons.nal.usda.gov/articles/dataset/USDA\_Table\_of\_Nutrient\_Retention\_Factors\_Release\_6\_2007\_/24660888 The USDA nutrient retention factors database has information on % retention, based on food types, for micronutrients for a range of cooking methods such as baking, broiling, and boiling. For the calculation of nutrient retention for INDB, the raw food ingredients were matched to the most commonly practiced cooking methods in Indian recipes.

**There are 2 information files:**

1. Excel file called ‘recipe\_links.xlsx’

This contains the name, food code, and web link (e.g., URL) of the recipes (n=150) that were taken from open online sources, thus the recipes were coded with the prefix Open Source Recipes (OSR).

2. Excel file called ‘units.xlsx’

This contains information on unit conversions that were used for entering the amount of raw ingredients for each recipe. The amounts of the raw ingredients for each recipe were entered as grams, millilitres, tablespoons, teaspoons, or sprigs (the dataset mentioned in #4). In some cases, the amounts were entered as numbers such as one apple. In these cases, the information on ‘weights and equivalent measures of some food stuffs’ in both the ‘Basic Food Preparation, 4th Edition’ and ‘The Art of Science of Cooking, 5th Edition’ manuals were used to convert the amount of the food item to grams, tablespoons, and teaspoons. In some cases, the USDA food composition database is used to obtain the data for unit conversions.

**There is 1 code file:**

1. Stata file called 'INDB.do'

This uses the FCT datasets described above to derive the nutrients per 100g recipe and per serving size of recipe. It then appends the recipe nutrient information to the ingredient nutrient information for a master IFCT with both foods and recipes.